AL-TR-1992-0083

AD-A262 920

ARMSTRONG LABORATORY

THEORY OF SIGNAL DETECTION AND ITS APPLICATION TO VISUAL TARGET ACQUISITION: A REVIEW OF THE LITERATURE

Denise L. Wilson

CREW SYSTEMS DIRECTORATE
HUMAN ENGINEERING DIVISION

MAY 1992

FINAL REPORT FOR PERIOD OCTOBER 1991 TO MAY 1992

Approved for public release; distribution is unlimited.

AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Armstrong Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AL-TR-1992-0083

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Kenneth R Boff

KENNETH R. BOFF, Chief
Human Engineering Division
Armstrong Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
May 1992

3. REPORT TYPE AND DATES COVERED
Final Oct 1991 - May 1992

4. TITLE AND SUBTITLE
Theory of Signal Detection and its Application to Visual Target Acquisition

5. FUNDING NUMBERS
PE - 62202F
PR - 7184
TA - 10
WU - 46

6. AUTHOR(S)
Denise L. Wilson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Laboratory, Crew Systems Directorate
Human Engineering Division, HSD/AFMC
Wright-Patterson AFB OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER
AL-TR-1992-0083

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

An operator's performance in a target acquisition task may be influenced by variables which are not directly associated with the target and imagery but are associated with the decision rule employed by the operator. The separation of factors that influence the operator's decision criterion from those that influence sensitivity is a major contribution of the psychophysical application of the Theory of Signal Detection (TSD). This report focuses on practical application of TSD to the problem of operator acquisition of targets in sensor imagery which may, or may not, have been prescreened by an automatic target recognizer. The foundation of TSD in Statistical Decision Theory is discussed as well as specific methods for applying TSD to laboratory evaluation of the target acquisition task. A brief review is presented of target acquisition studies in which TSD was utilized.

14. SUBJECT TERMS
Target Acquisition, Signal Detection
Decision Theory

15. NUMBER OF PAGES
56

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

PREFACE

This work was performed under Air Force Work Unit 7184 10 46, "Theater Battle Management." It was performed by the Crew Station Integration Branch, Human Engineering Division, Crew Systems Directorate, of the Armstrong Laboratory.

The effort resulted from the author's participation in the Automatic Target Recognizer Working Group (ATRWG). Special acknowledgements are due to Mr Lloyd Clark and Mr Raymond Wabler, respectively, the Department of Defense (DoD) and Industry co-chairmen of the ATRWG Applications Panel, and to Mr Gilbert Kuperman, the DoD co-chairman of the ATRWG ATR/Man-Machine Interface Panel, for their support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

Section 1

INTRODUCTION

HISTORICAL PERSPECTIVE

The Theory of Signal Detection (TSD) began with theoretical developments made in electrical engineering in the early 1950's. Mathematicians and engineers saw relevance of statistical decision theory to the general detection problem. They combined decision theory with an elaboration of the concept of the ideal observer, from electronic communications and radar theory, to form a general theory of signal detectability (Swets, 1964).

The part taken from statistical decision theory provides a way of controlling and measuring the criterion the observer uses in making decisions about signal existence, thus providing a measure of the observer's sensitivity that is independent of his decision criterion. This measure of sensitivity has been found to be practically invariant over several different psychophysical procedures, or detection tasks (Green and Swets, 1966).

The part that grew from work in electronic communications specifies the mathematically ideal detector and, therefore, ideal sensitivity, as a function of measurable parameters of the signals and the interfering noise, for several kinds of signals. This theoretical structure defines relevant physical variables in quantitative terms. It also describes the effects of changes in these variables, and how uncertainty about the specific values of these variables will affect certain detectors. Normative standards of performance allow one to gauge how the efficiency of human sensory processes varies with changes in the stimulus. This has allowed researchers to make inferences about the kind and extent of sensory information that is utilized by the observer, and about the nature of sensory processing (Green and Swets, 1966).

The theory specifies the mathematically ideal or optimal decision process. It is not intended to apply to any realizable sensing device, and was constructed without regard for human sensory processes. However, it soon became apparent that the general theory is a good approximation to a descriptive theory of human detection and recognition behavior. Researchers of psychophysics saw several analogies between this description of ideal behavior and various aspects of the perceptual process. Detection theory seemed to provide a framework for a realistic description of the behavior of the human observer in a variety of perceptual tasks (Swets, Tanner, and Birdsall, 1964).

The theory also serves as a guide for the study of human perceptual processes specifying appropriate experimental methods. The theory not only discloses new problems but also provides new approaches to old problems (Swets, 1964).

An experiment in vision (Tanner and Swets, 1954) was among earliest applications of detection theory in psychology. Smith and Wilson (1953) demonstrated its applicability to audition. Both of the above experiments concentrated on decision aspects of detection.

IMPORTANCE OF TSD FOR TARGET ACQUISITION

More than sensory information is involved in detection. The process of perceiving, as represented in Figure 1, is not merely one of passively reflecting events in the environment, but one to which perceiver himself makes a substantial contribution. The observer relates his sense data to information he has previously acquired, and to his goals, in a manner specified by statistical decision theory (Swets, Tanner and Birdsall, 1964). Most non-sensory factors are integrated into a single variable, the criterion. This results in a pure measure of sensitivity, largely unaffected by other than physical variables.

In classical psychophysical experiments, results expressed as thresholds were a function of both stimulus detectability and the observer's criterion. The threshold measure of sensitivity to stimuli may be contaminated by changes in the observer's criterion. Classic methods of psychophysics make effective provision for only a single free parameter, one that is associated with the sensitivity of the observer. They contain no analytical procedure for specifying independently the observer's criterion. These two aspects of performance are confounded in experiments in which the dependent variable is the intensity of the stimulus that is required for a threshold response. Changes in criterion, or different observers having different criteria will result in inconsistent responses to the same level of stimulus intensity within an individual observer or between observers. The application of TSD solves the problem in psychophysical

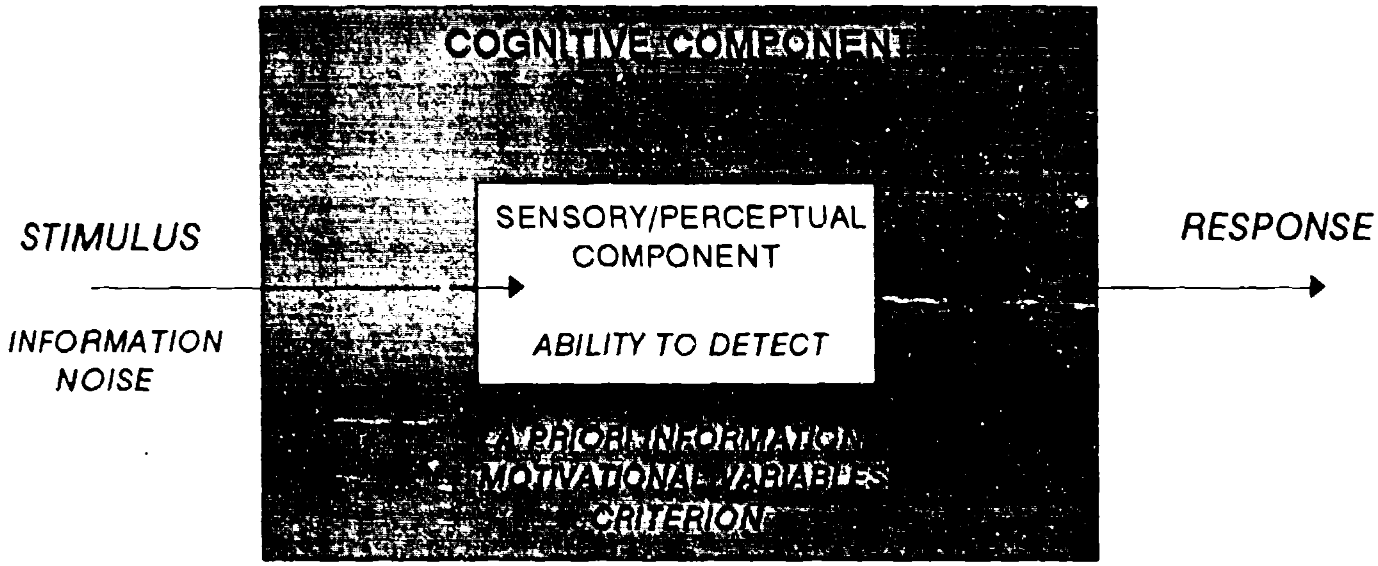

Figure 1. The process of perceiving.

experimentation of controlling or specifying observer criterion for making a perceptual judgement. TSD provides a quantitative measure of the criterion, leaving a relatively pure measure of sensitivity (Swets et al., 1966). TSD and associated methodology afford a means of independently measuring each of these factors.

Separation of factors that influence the observer's attitudes from those that influence sensitivity is a major contribution of the psychophysical application of statistical decision theory. This separation allows the specification of the perceiver's contribution to perception at other than conversational level, thus providing quantitative relationships between the nonsensory factors and both independent and dependent variables (Swets et al., 1964).

TSD is a normative theory. Having a standard with which to compare observer performance aids in description and interpretation of experimental results. (Swets et al., 1964).

SUMMARY OF REPORT

Although TSD has been applied to a number of areas of psychophysical research, this report focuses on the review of that literature which addresses its application to visual target detection/recognition. The focus of this report is heavily application oriented.

Part 2 presents a discussion of statistical decision theory, including the elements of the decision problem, assumptions, mathematical notation, decision outcomes, optimization of decision making, and decision goals.

Part 3 presents a discussion of TSD as applied to operator target acquisition performance. This includes a description of the properties of the target and background and discussion of the operator's decision rule. The separation of measures of operator perceptual sensitivity and response bias is discussed. Procedures for calculation of the d' measure of sensitivity and Beta (ß) criterion measure are covered. The theoretical significance of the ROC curve is discussed as well as interpretation of curves.

Part 4 discusses three basic procedures for data collection under the TSD paradigm. These include the yes - no procedure, the confidence interval method, and the forced choice procedure.

Part 5 gives a discussion of the underlying assumptions of TSD. This section also persents a discussion of appropriate tests to determine whether the assumptions have been met and application of non-parametric analysis methods to cases where specific assumptions have not been met.

Part 6 presents a discussion of the application of TSD to acquisition of targets in sensor imagery. Methodology and results of applicable studies are discussed.

Part 7 presents a summary evaluation of applicability of TSD to the visual target acquisition task.

Section 2

ELEMENTS OF THE THEORY OF SIGNAL DETECTION

STATISTICAL DECISION THEORY

The application of the theory of decision making to situations in which certain "signals" may or may not be added to a random background disturbance called "noise" is a major part of decision theory.

Decision Problem Minimal Elements

Within the context of the signal detection decision problem there are only two possible states of the world. These are the absence of a signal or the presence of a signal.

Another essential element of the decision problem is the information. This may be defined as the interval during which the observer attends to some sensory display or "makes an observation."

The final element of the decision problem is the decision itself. There are two possible decisions: (1) Signal Absent or (2) Signal Present (Green and Swets, 1966).

Assumptions

There are a number of assumptions associated with the decision problem. First among these is the assumption that the observer is fallible.

There are assumed to be different distributions of observations for the "signal" state of the world versus the "no signal" state of the world. This means that there exist differing probability densities for each possible observation given both the "signal" and "no signal" distributions. An observation is like an event in probability theory. That is, an observation is an element of a set such that a probability can be defined for each event. The probability of one of the events occurring on any given trial is unity.

It is assumed that the available responses are determined by the nature of the decision task. The available responses are often in one-to-one correspondence with the possible states of the world.

It is further assumed that the response of the observer is dictated by some policy or strategy on the part of the observer. This policy or strategy is referred to as the "decision rule." The "decision rule" is a mathematical function that maps the space of observations onto the space of responses. The "decision rule" can be expressed as a function, or a simple partition, of the elements of the observation class.

Notation

Most texts on statistical decision theory employ a similar notation. The notation described herein follows that of Green and Swets (1988).

The information or evidence, usually fallible, about a possible state of the world, h_i, is designated as e_k. The decision-maker selects an alternative state of the world, H_j, which may or may not be correct. There must be at least two possible states of the world, h_1 or h_2; h_j denotes the world in the jth state. The response of the decision-maker is then to accept one or the other hypothesis about possible states of the world; H_j denotes the decision-maker selection the jth hypothesis. The decision-maker's behavior is described as follows: $P(H_i|h_j)$ or "the probability of the decision-maker accepting the ith hypothesis about the state of the world given that the world is in the jth state." If $i = j$, the decision-maker's hypothesis about the state of the world matches the actual state of the world, and the response is correct.

The probability of occurrence of any single e_k usually depends upon the state of the world. $P(e_k|h_i)$ denotes the probability of e_k occurring given that the world is in the ith state. That is, the probability of e_k is conditional on h_i. The conditional probability, $P(e_k|h_i)$, is generally not equal to the conditional probability, $P(e_k|h_j)$.

An a priori probability is the probability of any particular state of the world prior to the observation. $P(h_j)$ represents the a priori probability of any particular hypothesis being true.

A posteriori probability is the probability of the hypothesis conditional on the occurrence of the observation. This is the probability of the truth of each hypothesis after the event has occurred. $P(h_i|e_k)$ denotes the a posteriori probability of h_i being true, given that event e_k has occurred. The relationship of a posteriori probability to conditional and a priori probabilities is as follows: $P(h_i e_k) = P(e_k)\ P(h_i|e_k)$ where $P(h_i|e_k)$ denotes the probability of the joint occurrence of h_i and e_k. $P(h_i e_k) = P(h_i)\ P(e_k|h_i)$. Since e_k can occur only if some hypothesis holds, $P(e_k) = \Sigma P(h_i)\ P(e_k|h_i)$. This leads to Bayes rule: $P(h_i|e_k) = [P(h_i)\ P(e_k|h_i)]/P(e_k) = [P(h_i)\ P(e_k|h_i)]/[\Sigma P(h_i)\ P(e_k|h_i)]$, which expresses the (<u>a posteriori</u>) probability that the world is in the ith state given that information e_k has been received (Skolnik, 1980).

The likelihood ratio is the ratio of the probabilities of an event under two different hypotheses. This ratio evaluates the evidence provided by an observation independent of the <u>a priori</u> probabilities of the hypotheses. It is a quantity that expresses the strength of evidence associated with each observation. For two hypotheses, $l_{ij}(e_k) = P(e_k|h_j) = 1/l_{ij}(e_k)$ where l_{ij} is read "the likelihood ratio of event e_k for hypothesis i relative to hypothesis j." If more than two hypotheses are involved in the decision problem, the number of likelihood ratios required to specify the decision problem completely is one less than the number of hypotheses. Likelihood ratio remains a single real number no matter what the structure of the observation. The observation may be multidimensional such as a certain configuration of weather -- temperature, wind velocity, time of year and other factors.

Decision Outcomes

Any single decision in a target detection task can be categorized as either right or wrong. An evaluation of performance is then generally based on some function of the average number of correct and incorrect decisions. There are always at least two types of correct decisions and two types of errors, as shown in the decision space depicted in Figure 2. The two types of correct decisions may be referred to as hit (declaration of a target detection when a target was indeed present) and correct rejection (declaration of no target detection when indeed no target was present). The two types of errors are false alarms (declaration of a target detection when in reality no target was present) and miss (declaration of no target present when a target was indeed present).

As an operator's decision criterion is shifted, it is possible to increase the percentage of hits or to decrease the percentage of false alarms. However, it is not possible to simultaneously increase the probability of both types of correct decisions. A decrease in the probability of one type of error is accompanied by an increase in the probability of the other type of error. That is, the operator may adopt a more stringent criterion, thus reducing false alarms, but would then increase the number of misses. If the operator adopts a more lax criterion, thus reducing the number of misses, the number of false alarms will then increase. Depending on the situation, some types of errors may or may not be more unacceptable than others. For example, the penalty may be greater for missed targets than for false alarms or vice versa.

In applying likelihood ratio to optimize decision making,

		TRUE TARGET CONDITION	
		TARGET	NO TARGET
REPORTED TARGET CONDITION	TARGET	*HIT*	*FALSE ALARM*
	NO TARGET	*MISS*	*CORRECT REJECTION*

Figure 2. Decision space consisting of two types of correct decisions and two types of errors.

the decision-maker may attempt to satisfy any of several goals. These include maximization of a weighted combination, maximization of expected value, maximization of percentage of correct responses, or satisfaction of the Neyman-Pearson objective. A decision rule that maximizes the weighted combination, $P(H_1|h_1) - \beta P(H_1|h_0)$ is to choose H_1 if and only if $l_{10}(e_i)$ is $\geq \beta$. The expected value is maximized by accepting as h_1 all those events whose likelihood ratio of h_1 to h_0 is equal to or greater than β as defined by:

$$\beta = (V_{00} + V_{01})P(h_0)/(V_{11} + V_{10})P(h_1),$$

where V_{00} = value associated with a correct choice of H_0,

V_{01} = value (cost) associated with incorrect choice of H_0 when, in fact, H_1 is the correct alternative

V_{11} = value associated with a correct choice of H_1

V_{10} = value (cost) associated with an incorrect choice of H_0 when, in fact, H_1 is the correct alternative.

If all values associated with correct decisions are equally valued and all errors are equally intolerable, then the goal may be simply to maximize the percentage of correct responses. That is, the value of being correct is set at unity and the cost of an error is et at zero. The percentage of correct decisions is maximized if $\beta = P(h_0)/P(h_1)$. The Neyman-Pearson objective is as follows: for some constant k, where $0 \leq k \leq 1$, set $P(H_1|h_0) = k$ and then maximize $P(H_1|h_1)$. This goal is utilized by those who perform statistical tests. k ia usually set to 0.01 or 0.05 and then one endeavors to design a test to maximize the acceptance of

h_1 when it is true while holding the error rate at the level k. The decision rule is therefore: Choose H_1 for all events whose likelihood ratio is equal to or exceeds ß, where ß is selected such that $P(H_1|h_0) = k$ (Green and Swets, 1966).

SUMMARY

Within the context of the signal detection decision problem, there are two possible states of the world and two possible decisions. This results in four possible outcomes, two of which are correct decisions and two of which are incorrect decisions. It is assumed that the response of the observer is dictated by some policy or strategy which is referred to as the "decision rule." Any of a number of "decision rules" may be employed. Decision rules are often related to the likelihood ratio; that is, the rule states that H_1 should be chosen as a response if $l_{10}(e_i)$ is $\geq$ ß, where ß is some critical value corresponding to the goal of the decision-maker.

Section 3

APPLICATION OF THEORY OF SIGNAL DETECTION TO VISUAL TARGET ACQUISITION

An operator's performance in a target detection task may be a function of (1) properties of the target and imagery such as image resolution, grazing angle, degree of camouflage, etc.; and (2) operator's decision rule (e.g., presumed to be related to the payoff for correctly identifying a target and the consequences associated with false alarms). Performance measures such as probability of hits and probability of false alarms do not allow one to unambiguously interpret the results; that is, the contribution to the response measures associated with stimulus characteristics and response bias are not separable. For example, an operator may achieve a high probability of hits, but achieve this result by calling everything a target (e.g., a high probability of false alarms).

The Theory of Signal Detection (TSD) provides a means by which one can obtain two independent measures that relate to operator sensitivity and response bias, respectively (Green and Swets, 1966). The measure of sensitivity, referred to as d', is generally affected by sensory/perceptual factors such as image resolution and target-to-background relationship. ß, on the other hand, is a measure of response bias which is affected by such variables as the consequences of misses and false alarms, rules of engagement, _a priori_ knowledge, expectations, and training. The value of β is an index of the operator's response criterion.

The operator must establish a policy that defines the circumstances under which the observation will be regarded as resulting from each of the two possible events. Since operator is assumed to be capable of locating a criterion at any point along the continuum of observations, it is of interest to examine the various factors that, according to the theory, will influence his choice of a particular criterion. (Swets et al., 1964)

In a signal detection experiment the following performance parameters are computed:

(1) Probability of a hit [p(Hit)]

(2) Probability of a false alarm [p(FA)]

(3) Probability of a miss [p(Miss)]

(4) Probability of a correct rejection [p(CR)]

Figure 3 illustrates how these performance parameters are related to the noise and signal-plus-noise distributions.

The theory proposes that d' is equal to the difference between the means of the signal and noise (SN) and noise (N) distributions ($u_{SN} - u_N$) expressed in standard deviation units of the N distribution.

Because the location of the SN distribution with respect to that of the N distribution is entirely a function of stimulus intensity and properties of the sensory system, d' is a pure index of stimulus detectability which is independent of the location of the operator's criterion (β).

The value of ß simply weights the hits and false alarms. ß is determined by the _a priori_ probabilities of occurrence of signal and of noise alone and by the values associated with the individual decision outcomes. Optimal cutoff along the x axis is at the point on this axis where the ratio of the ordinate value of the signal and noise function to the ordinate value of the noise function is a certain number, ß. ß specifies the optimal weighting of hits relative to false alarms. The operator's criterion should always be located at the point on the x axis corresponding to the value of ß. For any detection goal to which the operator may subscribe, and for any set of parameters that may characterize a detection situation (such as _a priori_ probabilities and values associated with decision outcomes) the optimal criterion may be specified in terms of a single number, ß (Swets _et al_., 1964).

In signal detection analysis, the corresponding Z scores from a normal distribution for the proportions of hits and FAs are used to calculate d', which is the measure of target detectability or operator sensitivity (Gescheider, 1976). A Z-score presents the number or standard deviation units that a particular hit rate or false alarm rate is from the mean of a standard (mean = 0, standard deviation = 1.0) normal distribution. The equation used to calculate d' is:

$$d' = Z(\text{Hits}) - Z(\text{FA})$$

ß, on the other hand, is the ratio of the ordinate of the SN distribution at the criterion to the ordinate of the N distribution at the criterion, as follows:

$$\beta = \frac{f_{SN(X)} \text{ at criterion}}{f_{N(X)} \text{ at criterion}}$$

where f is the function for computing the ordinate of X along the normal curve. A low value of ß represents a lax criterion where the operator will be liberal about reporting "signals," while a high value of ß represents a strict criterion where the operator will be conservative about reporting signals.

Figure 3 shows the relationship between p(Hit), p(FA), p(Miss) and p(CR) for two identified curves representing the N and SN distributions. The criterion line in this Figure is just one of an infinite number of criterion lines. The setting of the criterion line affects the value of ß and is independent of d'. If the criterion line were shifted to the left, the value of ß would decrease, indicating that the operator is using a lax criterion. On the other hand, if the criterion line were shifted to the right, the value of ß would increase, indicating that the operator is using a strict criterion. Movement of the criterion line, and hence the value of ß, has no effect on d' since this value is a function of the distance between the SN and N distributions. The distance between these two distributions affects the operator's sensitivity in detecting signals from the noise.

In TSD studies, both parameters affecting perceptual sensitivity and parameters affecting decision criterion may be manipulated. This allows the determination of receiver operating

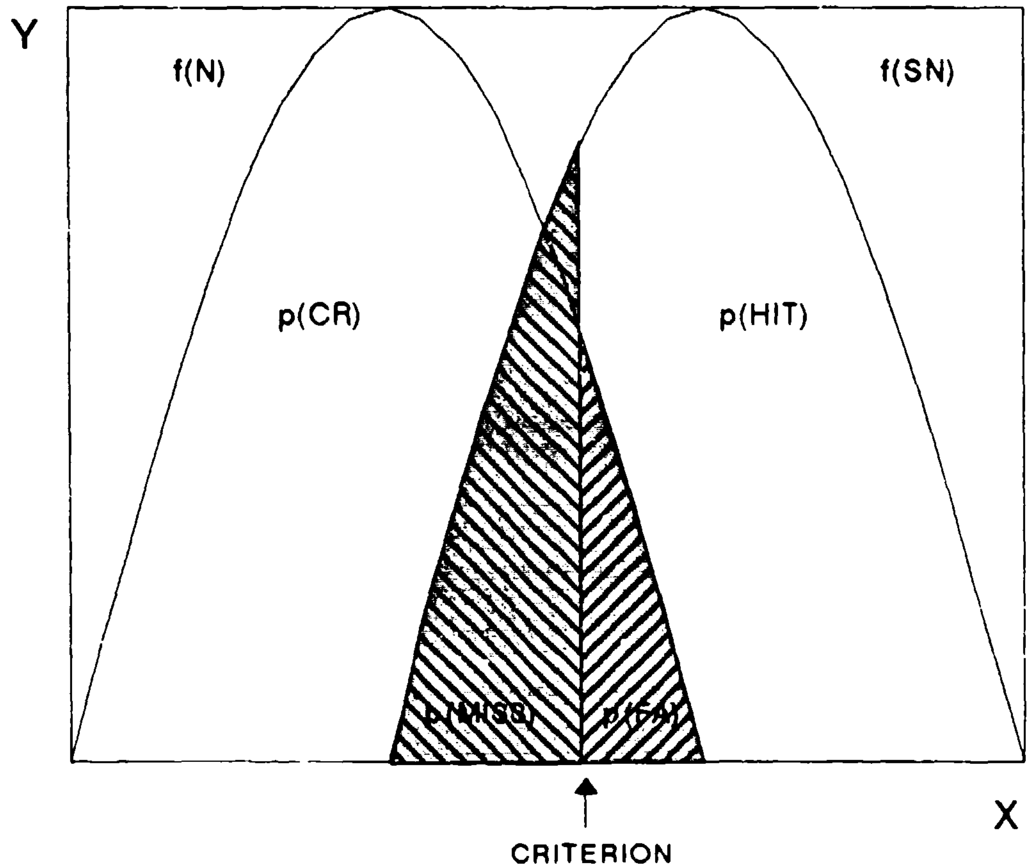

Figure 3. Relationship of p(Hit), p(FA), p(Miss) and p(CR) to N and SN distributions.

characteristic (ROC) curves (Gesheider, 1976). An ROC shows the relationship between false alarm and hit rates. Figure 4 presents a family of ROC curves with d' values ranging from 0.0 (chance performance) to 3.0 (near-perfect performance). Points along a given ROC are generated as a function of changes in the operator's criterion. The value of can be computed from an ROC by determining the slope of a tangent at any given point along the ROC.

Figure 5 shows the effect of changes in ß on the shape of the ROC. As can be seen, variations in the operator's criterion results in different points along the same ROC curve while variation of signal strength produces different ROC curves.

The ROC curves are useful in evaluating whether the assumptions underlying the TSD paradigm have been met. In addition, they are useful in determining if the effects of an experimental variable on performance are due to changes in response criteria. (e.g., overlapping ROC curves) or to differences in sensitivity (e.g., different ROC curves that do not overlap).

In order to generate ROC curves, pairs of H(hit) and p(FA) must be generated where the operator's response criteria varies. This can be accomplished by varying such variables as: (1) target probability, and (2) payoffs and consequences of hits and false alarms. These variables would affect an operator's response criteria and therefore would yield different points in an ROC space such that a curve characterizing performance could be drawn.

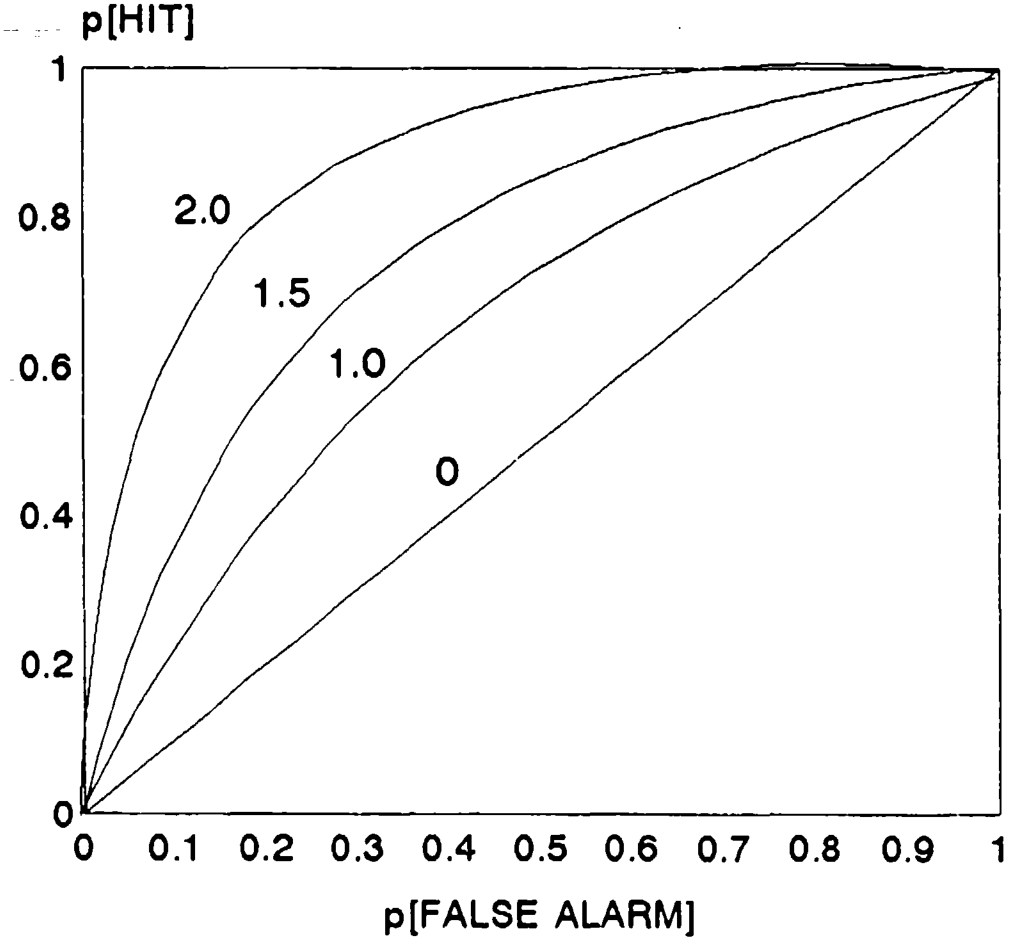

Figure 4. Family of ROC curves with d' values ranging from 0.0 to 2.0.

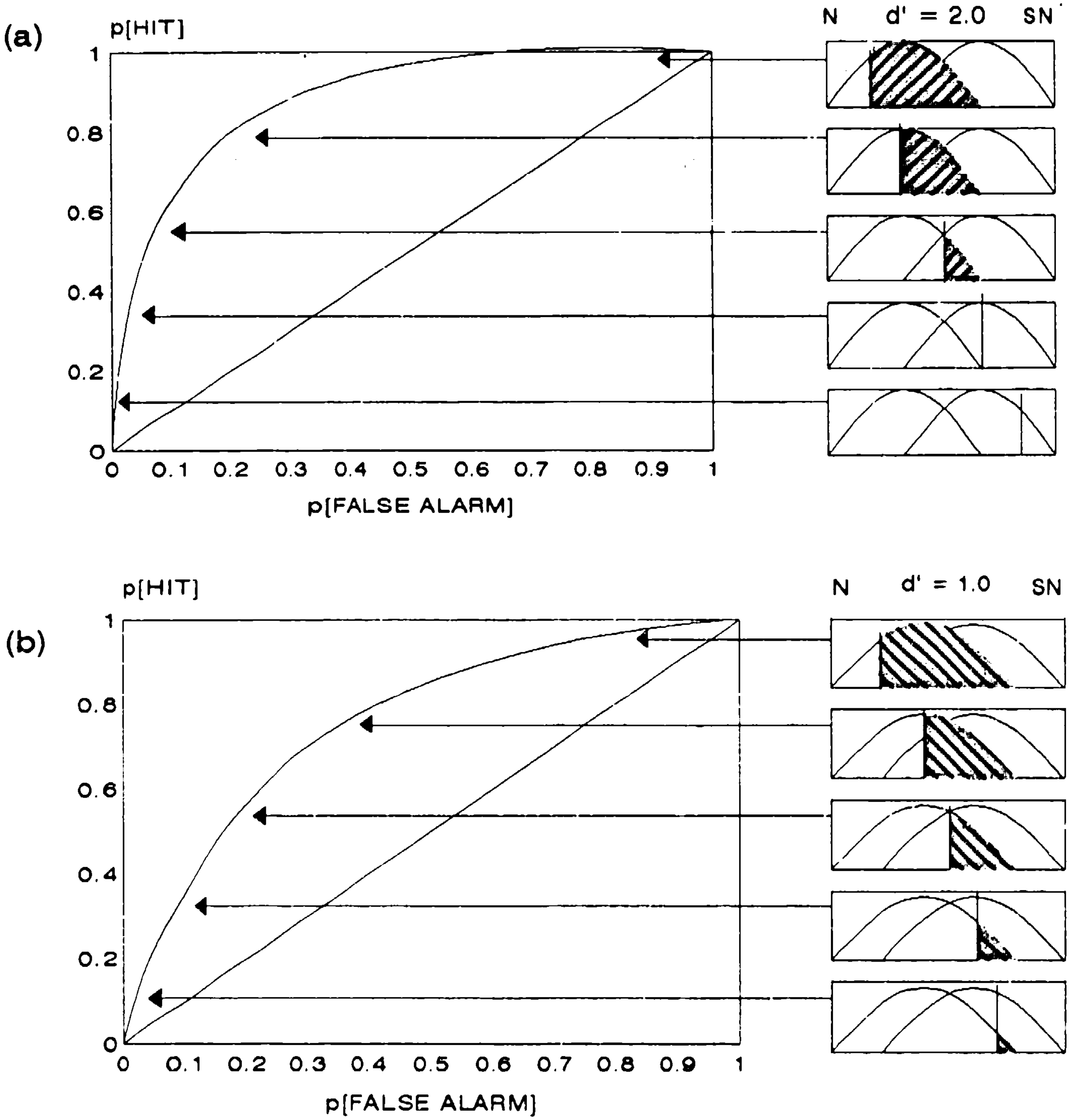

Figure 5. Illustration of the manner in which ROC curves are predicted from TSD. (a) depicts a situation where signal strength is sufficient to result in only a slight overlap of the N and SN distributions, while (b) depicts a situation where signal strength is weak, resulting in considerable overlap.

One of main sources of evidence supporting TSD is experimental manipulation of variables resulting in data plotted as ROC curves. Shapes of ROC curves for various stimulus intensities can be generated from the postulates of the theory and checked against empirical data. Variation in signal strength produces different ROC curves. Variation in the operator's criterion results in different points along the same ROC curve (Gescheider, 1976).

To illustrate the manner in which ROC curves are predicted from TSD, Figure 5 depicts a situation where signal strength is sufficient to result in only slight overlap of the N and SN probability distributions. The vertical lines represent locations of criterion that might be associated with particular conditions of stimulus probability and payoffs. According to TSD, each point on an ROC curve is determined by the location of the operator's criterion. If the observation is to the right of the criterion, the operator will say "yes." The proportion of yes decisions is equivalent to the proportion of the area under the curve to the right of the criterion. Values of p(yes/N) and p(yes/SN) are determined by finding the areas under the N and SN curves, respectively, which are to the right of the criterion. As the criterion is changed, the values of the p(yes/N) and p(yes/SN) change. These values, when plotted, form an ROC curve.

The lower curve of Figure 5 depicts the situation where signal strength is so weak that it results in considerable overlap between the N and SN distributions. TSD is strongly supported by the finding that theoretical ROC curves generated in this manner are very similar to those obtained experimentally (Gescheider, 1976).

The theoretical concept of signal detectability can be measured by determining on which member of a family of ROC curves an operator's responses fall, thus ascertaining the approximate value of d'. Exact values of d' can be derived from the empirical values of p(yes/SN) and p(yes/N). For a particular separation of the SN and N distributions, the value of d' will remain constant for all possible criterion positions. The ROC curve is, therefore, a description of performance changes which are accounted for by a constant d' and a continuously variable criterion (Gescheider, 1976).

Once the correct ROC curve has been determined, the location of the operator's criterion, ß, can be determined by observing exactly where on the ROC curve the point is located. If the point is near the bottom of the ROC curve, where the slope is great, the criterion is high; if the point is near the top of the curve where the slope is slight, the criterion is low. The exact value of ß is equal to the slope of the ROC curve at a particular point. ß is a value of the likelihood ratio; the ratio of the ordinate of the SN distribution at the criterion to the ordinate of the N distribution at the criterion.

Section 4

THREE PROCEDURES USED IN TSD

One of three procedures is generally employed in TSD studies. These are the yes - no procedure, the forced choice procedure, and the confidence rating procedure.

YES-NO PROCEDURE

Observers are given a long series of trials (usually more than 300 per session), some proportion of which are SN while the others are N only. For each observation, one of two mutually exclusive alternatives is presented within a clearly marked observation interval. The observer is usually told what proportion of the trials will contain a signal as well as the payoffs and penalties associated with the four possible decision outcomes. The observer is asked to respond by selecting one of the two permissible response alternatives (Gescheider, 1976; Green and Swets, 1966).

An ROC curve can be plotted for a single signal strength if the proportions of hits and false alarms are plotted for several criterion locations. Data for different criterion levels are often obtained by changing signal probability or payoff contingencies for different sessions. Tanner, Swets, and Green (1956) have shown experimentally that while large variations in the observer's criterion were produced by two distinct procedures (variation of signal probability and variation of payoff contingencies) the d' measurement of sensitivity remained stable under all conditions. This is as predicted by TSD, because the physical stimulus remained the same and the shape and locations of the N and SN distributions did not change. Therefore sensitivity did not change (Gescheider, 1976).

If data are collected for a single session, and proportions of hits and false alarms are available for only one criterion location, the values of d' and ß can be estimated from the data. However, an ROC curve cannot be plotted. The difference between the z scores for hits and false alarms will yield an estimate of d'. The value of ß can be obtained by dividing the ordinate value on the normal curve corresponding to the z score for hits by the ordinate value corresponding to the false alarms. When ROC curves are not available to check the validity of the normal distribution and equal variance assumptions, measures of sensitivity not requiring these assumptions should be used whenever possible (Gescheider, 1976). Such a nonparametric measure of sensitivity, termed A', has been proposed by Pollack and Norman (1964). The formula for calculating A' is:.1s1

$$A' = \frac{1/2 + [p(\text{hits})-p(\text{false alarms})][1-p(\text{hits})-p(\text{false alarms})]}{[4p(\text{hits})]\ [1-p(\text{false alarms})]}$$

(Gescheider (1976).

FORCED CHOICE PROCEDURE

The forced choice procedure is an excellent technique for obtaining a measure of the observer's sensitivity which is uncontaminated by fluctuations in his criterion. On a particular trial, two or more observation intervals are presented and it is the observer's task to report which observation interval contained a signal. The assumption is made that in the absence of response bias toward one or more of the observation intervals the observer chooses the observation interval containing the largest sensory observation. Since the observer's criterion is not a factor in such a judgement the proportion of correct responses p(c) can be used as a measure of sensitivity. The value of p(c) will be underestimated when response bias toward one of the observation intervals exists. Procedures for correcting the p(c) obtained when response bias exists are found in Green and Swets (1966).

CONFIDENCE RATING PROCEDURE

The confidence rating method allows the experimenter to obtain an ROC curve from data in a single session within which signal probability and payoff contingencies are fixed. The method is very economical; data for several points on an ROC curve can be obtained for a single experimental condition by having the observer make a confidence rating for each of his yes - no judgements. It is assumed that to make his ratings, the observer sets up n-1 criteria along the sensory continuum to delineate his rating categories. The number of criteria is one less than the number of rating categories (Gescheider, 1976).

During the experimental session, the proportion of responses for each of the rating categories for the SN trials and for the N trials are determined. One can then calculate the hit and false alarm rates that would occur if the observer were induced to set his yes - no criterion at each of the n-1 criterion points defined by the n rating categories. Each of the n-1 pairs of hit and false alarm proportions that result from this procedure provides a point on an ROC curve (Gescheider, 1976).

Because sufficient data can be quickly obtained for constructing an ROC curve by the confidence rating procedure, its use can provide a convenient means of testing the hypothesis of normality of N and SN distributions and the equal variance of N and SN distributions. Green and Swets (1966) have demonstrated the validity of the confidence rating procedure. The confidence rating procedure was found to yield very similar values of signal detectability to those obtained by the yes - no procedure (Gescheider, 1976).

Section 5

ASSUMPTIONS ASSOCIATED WITH TSD AND THEIR TESTS

The form of the ROC curve predicted from TSD can be more easily subjected to experimental tests if the values of p(yes|SN) and p(yes|N) obtained in an experiment are plotted on the ROC curve as z scores. If the N and SN distributions are normal in form and also have equal variances, the ROC curves should be linear with a slope of 1.0 when z scores for hits are plotted against z scores for false alarms. In the normal distribution and equal variance situation, when the criterion is shifted by a particular z score distance on the N distribution, it is also shifted by exactly the same distance on the SN distribution. The linearity prediction follows from the assumption that the N and SN distributions are normal in form. The prediction of a slope of 1.0 follows from the assumption of equal N and SN variances. The prediction from TSD is that ROC curves plotted from z scores should be linear with a slope of 1.0. Standard procedure is to determine the best fitting straight line for the data plotted as z scores. The use of the method of least squares will provide the best estimate of the intercept and slope of the function. If the data points do not significantly deviate from the function, the assumption of normal distribution is supported. If the slope of the function does not significantly deviate from 1.0, the equal variance assumption is supported.

Empirical ROC curves plotted as z scores are almost always linear, leading to a rather general acceptance of the hypothesis that the N and SN distributions are normal. The assumption that the N and SN distributions have equal variances is not supported in most experiments. The slope of the ROC curve is frequently found to be less than 1.0, a result usually explained by assuming that the variance is greater for the SN than for the N distribution. The variance of the SN distribution is assumed to increase as the mean of the distribution increases.

In cases where the variance of the SN distribution is greater than that of the N distribution, the symbol δm, rather than d', is sometimes used to denote the difference between the means of normal N and SN distributions. The quantities d' and δm are symbols for the same measures of signal detectability applied to the cases of equal and unequal variances, respectively. The value of δm is equal to the absolute difference between z(yes|N) and z(yes|SN) at a point where z(yes|SN) is equal to 0.

A measure of signal detectability that is sometimes used instead of δm is d_e'. The value of d_e' is the absolute dfference between z(yes|N) and z(yes|SN) at a point on the ROC curve where it crosses the negative diagonal. The primary benefit of using this measure is that it gives equal weight to s_N and s_{SN} (Gescheider, 1976).

Section 6

APPLICATIONS OF TSD TO VISUAL TARGET ACQUISITION TASKS

GREENING (1986) AND GREENING AND FOYLE (1988)

Greening (1986) investigated the applicability of the signal detection paradigm to the investigation of targeting decisions with multiple data sources. The context for this investigation was a Navy air-attack targeting system making use of data from on-board and remote imaging sensors, briefings and nonimaging sources. Greening defines the decision situation as extremely difficult due to requirement to make a difficult set of targeting decisions under conditions of extreme risk and urgency. He points out that there is uncertainty about the validity of every information source, and that difficult choices must be made concerning the employment of various sensors and data sources, and the interpretation of their outputs. The targeting function involves a search, the detection of a candidate object and its classification as target or nontarget. Greening points out that

> "As the number and variety of target data sources increase, the task of the observer changes character. Instead of the deep-seated and over-learned process of integrating and understanding a sequence of direct visual glimpses of the earth's surface, the observer must select the appropriate sensor(s), translate differing scales and coordinate systems, and interpret images, as well as nonimage data which will include nonvisual features of the target and environment. In time-limited scenarios especially, the potential utility of the multiple-source data may be more than offset by the complexity of the perceptual and decision process required of the observer." (pg 4)

Greening's analysis of the target detection problem takes an inside-out approach which can be characterized as starting with an aircrew member, embedded in an attack mission, faced with a targeting assignment. The focus of the study was to describe his targeting decisions and determine what information can be brought to bear on those decisions. One way of looking at the targeting decision process is that the observer evaluating a data input must establish a criterion "score" for inclusion of an object in the target class. If his criterion is set too high, he will almost surely be right when he classifies an object as a target, but will miss many other targets. Too low a criterion will include more targets but also more false alarms.

The theory of signal detection provides a means of connecting performance to sensor quality and to the kinds of confusion objects present (through the sensitivity parameter, d'). The receiver operating characteristic (ROC) concept and the ROC plot provides a means of trading off criterion level and classification performance. The extension of signal detection methods to multiple observations with adjustable criterion levels provides a means for estimating the performance gains which may be made using successive observations with the same sensor or a different sensor. Decision theory provides formal procedures for optimizing decisions, based upon assigned costs and values and known or estimated probabilities of success. While quantification of costs and values in combat situations is difficult, the theory will show the sensitivity of outcomes to postulated changes

(Greening, 1986).

Greening (1986) describes an imaginary test situation in which an observer is shown a long series of images. After each observation, the observer must make a judgement as to whether the object is a member of a target class (and how sure he is). The judgement might be expressed as a rating, varying from "certain target" to "certain nontarget" with any number of intermediate levels of uncertainty.

In general, the observer doesn't know whether the judgement is right or not. The images come from two distributions (target and nontarget) which are likely to overlap to some extent (i.e., some target images are easily confused with nontarget images).

The observer is faced with a decision related to the observed image. For an attack mission, he must decide whether a particular image justifies releasing a weapon. He must decide what to do with the equivocal judgements, without knowledge of the underlying target/nontarget distributions. Depending on the tactical situation, he will decide to attack only if the score reaches some specific level (e.g., "probably a target" or higher). If the image in question meets this criterion, he will act as though the object in the image really is a target, even though he may be wrong some of the time. The outcome of the observation, then, will be a target declaration if it reaches or exceeds the criterion score, and a nontarget declaration if it fails to reach that criterion.

The joint occurrence or a true target and a target declaration will result in a hit, which is just one of the four possible outcomes. Other possible outcomes are false alarms, missed targets, and correct declaration of nontargets.

The hit and false alarm measures of performance are defined by TSD to be independent of target abundance. These two measures can be plotted against each other to give a composite indication of performance. If these values are plotted with false alarm probability on the abscissa and hit probability on the ordinate, they can be connected by a smooth curve which is the locus of all the pairs of values expected if the criterion was to move smoothly and continuously from left to right across the joint distributions of noise and noise plus target.

Greening (1986) cautions that because TSD was developed to describe and predict behavior in a comparatively simple, controlled laboratory setting, some thought must be given to the appropriate performance measures and to their relationships to the tactical situation. At one extreme one might consider the isolated attack aircraft with a sensor and an assigned target type. Each observation must be judged as a target or nontarget. Because he receives no vectoring information and no confirmation, the observer's best estimate of his performance may be the perceived target abundance. Thus, he may try to shift his criterion to bring his target declarations into line with the estimated target abundance. If he does this, the actual hit rate could drop to very low levels for many conditions.

If, on the other hand, confirmation of targeting decisions is available after the decisions are made, the attacker can be aware of hits, false alarms, missed targets, and rejected nontargets as they occur or soon after. From these data and the target density assumed previously, the observer may, in principle, tailor his criterion location to the tactical requirements. The major obstacle of using such a tuning procedure is the complexity of the data-management, calculation and optimizing procedures.

Greening (1986) and Greening and Foyle (1988) also explore the effects of multiple observations, whether these observations are independent or correlated. For the case of multiple independent observations, we might consider an aircraft with two different, but equally sensitive sensors, both observing the same object. The accuracy of classification might be expected to be improved over either sensor alone. For those cases in which the outcomes agree, the probability that the consensus is correct would be higher. The question is what is to be done for cases of disagreement. If these cases are simply discarded or treated as nontarget declarations, the overall performance level will drop. Greening proposes two approaches to avoiding the equivocal cases and thereby improving overall performance with two independent observations. The first approach is that of contingent criterion. Here the results of the first observation are used to determine the location of the criterion for the second. For example, if the first sensor gave a very high likelihood of a target, the criterion level for the second can be lowered, and vice versa. The other approach Greening terms optional stopping. Using this technique, the results of the first observation are used to determine whether more observations are necessary. If the first observation gives a substantially high or low score, stop observing; if the score is intermediate, take another observation. If still more observations are feasible, the process can continue, though the gains will fall off rapidly.

Probably a more common example of multiple observations in target acquisition will be successive looks with the same sensor. Here, the independent observation assumptions cannot be expected to hold. To the extent that successive observations are correlated, the amount of new information conveyed by the second observation will be reduced. In estimating the effect of multiple serial observations of the same object with the same sensor, we run into the problem of redundant information or observation-to-observation correlation. The correlation between responses on successive observations would be expected to be rather high. As a result, a second correlated observation might be expected to raise overall performance (as measured by d') very little.

Greening and Foyle (1988) discuss the limitations of statistical measures of targeting performance. A given target sensing system can be used to maximize hit rate or minimize false alarms, but can't improve all measures at once. The choice of criterion values must depend upon tactical considerations. This brings about the necessity of attaching a value, or cost, to each possible outcome of the decision problem. Greening and Foyle

(1988) suggest that the cost/value question be treated with the Expected Value model borrowed from economic analysis. The Expected Value model requires that a cost/value be placed on every possible outcome of an attack mission, and a cost placed on each observation. The likelihood of each outcome at each decision point must then be evaluated. Assuming that the observer is rational, the choice with the higher Expected Value can be assumed to be chosen at each choice if the Expected Values are available. The Expected Value of the entire mission, from start to finish, can then be calculated. The process is rather complex and the difficulty of obtaining meaningful, comparable costs and values for all outcomes should not be minimized. However, the Expected Value approach may suggest ways of viewing the problem which can be applied to real world target detection tasks. The Expected Value measure is attractive because it reflects all of the contingent probabilistic measures. However the task of obtaining meaningful value and cost figures is difficult, and the chosen costs/values are generally highly subjective.

KUPERMAN, WILSON AND PEREZ (1988)

Kuperman, Wilson and Perez (1988) conducted an operator performance study in which simulated synthetic aperture radar imagery containing either targets (large military vehicles) or nontarget distractors (smaller military vehicles) was employed in a target detection, recognition, and designation task. In addition, a no target/no distractor condition was included to support estimation of the frequency of occurrence of background-induced false alarms.

A signal detection paradigm was employed since insight into the decision-making criterion (accept/reject the presence of a target) was desired along with performance (accuracy and speed) estimates. The use of simulated imagery allowed for experimental control of the independent variables (target condition, resolution) and, therefore, avoided the problems that would have resulted from reliance on actual imagery collections.

The stimuli were simulated synthetic aperture radar images created via a radar model developed at the University of Kansas (Geaga, 1985). Four distinctly different backgrounds were created, essentially by means of varying the relative degree of forest coverage and tree height. Each of the four backgrounds was created as though imaged from three look directions (radar azimuths) in increments of 120 degrees.

Each of the resulting 12 scenes was generated under the following target conditions: (1) one target, (2) three targets, (3) one distractor element, (4) three distractor elements and (5) no targets or distractor elements. Two radar resolutions were simulated for each image. Thus, there were 120 unique images (four backgrounds by three look directions by five target/distractor/nontarget conditions by two resolutions).

For the lower resolution images, targets subtended approximately six minutes of arc with 8.2 lines on target, while distractor elements subtended approximately four minutes of arc with 5.4 lines. For the higher resolution images, targets subtended approximately eighteen minutes of arc with 24.7 lines on target, while distractor elements subtended approximately twelve minutes of arc with 16.1 lines. All images were scaled to the 0 to 255 range of intensity values.

The yes - no procedure was employed for the signal detection task. The subjects' task was to search the image for the target of interest. If no targets were present the subjects were to press the "no target" (reject) button. If one or more targets were thought to be present, the subjects' task was to move a video cursor (using a trackball) from the center of the display and place it over the center of the target and press a "designate" button. Search times were recorded as well as the "target" or "no target" declarations.

Data collection took place over five experimental sessions, each representing one replication of the entire set of 120 images. Order of presentation within each session for each subject was completely randomized with respect to all experimental variables.

The subject's response under each condition could be characterized as a hit, miss, false alarm, or correct rejection. The percentage of hits was viewed as an indication of the discriminability of the targets from the distractors as a function of background clutter and image resolution, whereas false alarms were a measure of distractor effectiveness.

An analysis of hit rate revealed that targets presented under the high resolution condition were responded to more quickly and more accurately than those presented under the low resolution viewing condition. False alarm rates were found to be affected by the number of distractors in the image and, to a lesser extent, by the image resolution. Response times also showed an effect due to number of distractors. Neither false alarm rate nor response times were greatly affected by resolution.

d' and ß were computed as a function of image resolution. The subjects demonstrated themselves to be more sensitive to the difference between target and background surround (e.g., clutter and distractors) under the high resolution viewing condition relative to low resolution images. The value of beta was lower under high resolution relative to low resolution viewing conditions. This indicates that subjects adopted a more stringent criterion under the low resolution viewing condition; that is, the subjects were more cautious when inspecting a low resolution image than when they were inspecting a corresponding high resolution image. Thus, the signal detection analysis procedure allowed the researchers to unambiguously determine if a variable affected perceptual processes, decision making processes or both.

OZKAPTAN (1979)

Ozkaptan (1979) applied the principles of the theory of signal detection to a target acquisition experiment which involved three types of instructions and two levels of target-to-background contrast. The purpose of the study was to evaluate the utility of the statistical parameters of the theory of signal detection when used to adjust the responses (reaction time and number of hits) of test participants to a comparable level of performance which is free of the effects of response bias that was introduced by the type of instructions.

Ozkaptan (1979) wished to develop a capability for transforming operationally relevant measures to a bias free level. The operationally relevant measures of interest were the number of hits and time to detection which are the usual dependent measures of a target acquisition experiment. The d' and ß values of a signal detection experiment, as dependent measures, are usually evaluated in and of themselves without comparison to other dependent variables. Ozkaptan felt that it was difficult to evaluate the implications of these measures relative to the operationally relevant dependent measures that may be associated with them. He felt that the direct evaluation of these operationally relevant measures, when free of response bias, would have more utility than information on d' and ß alone or when used in conjunction with them. Such bias-free measures would permit the direct comparability of data between experiments which differ in instruction and other utility variables.

A two factor experiment involving three levels of instruction and two levels of target-to-background contrast was employed. Twelve Army helicopter pilots were assigned to each instructional level, with target-to-background contrast as a within factor. The design was presented in the form of a 4 x 4

Latin Square to assure experimental control of the effects of trial sequence, target background, and the order of presentation of target-to-background contrast. The same design was repeated in a second experiment, using different test participants, in which only the levels of target-to-background contrast were changed. In the first experiment, contrast levels of 35 and 45 percent were used, which are typical of field studies, while in the second experiment, contrast levels typical of laboratory tests, 14 and 17 percent, were used.

A target acquisition task during a simulated helicopter pop-up maneuver at 1000 feet altitude was presented, with a 30-second exposure time. The observer's task was to search for a single 20-foot military tank in various field locations, and at a slant range of 2500 feet. An infrared scene of European terrain was simulated, which was presented on a 50 by 50 degree backlighted screen at 20 inches viewing distance. The scenes were presented with and without targets in order to obtain an observer's hit rate and false alarm rate, the basic procedural requirement of signal detection theory.

It was expected that target acquisition performance (reaction time and number of hits) would differ as a function of the type of instruction and level of target to background contrast presented to the test participants. It was further expected that the signal detection parameters of β and d' would reflect the effects of the different instructional levels and target-to-background contrast, respectively. The goal of the research was to evaluate the utility of the signal detection parameters of ß and d' when used in an analysis of covariance to adjust the operationally relevant dependent measures to a bias-free level of performance.

Analyses were conducted for eight dependent variables, which involved the combination of reaction time and frequency of response with respect to hits, false alarms, misses, and correct rejections. The signal detection parameters for each observer were calculated. The relationship between reaction time and the number of hits to the signal detection parameters were tested by means of regression analyses. The signal detection parameters were then used as covariates in a two-way analysis of covariance conducted separately for reaction time and the number of hits.

Ozkaptan (1979) found instructional set to be an important determinant of aviator performance during target acquisition with respect to reaction time and the number of hits. This effect may confound the results of similar experiments when this variable is left uncontrolled. He concluded that the signal detection model provides a reasonable representation of the sensitivity and bias effects associated with instructional set and target contrast, with some loss of precision due to its application under simulated "field" conditions. Further, the signal detection parameters associated with sensitivity and bias can be used in an analysis of covariance to adjust the frequency of hits between target acquisition studies, to remove the effects of different instructional sets as well as different contrast levels.

However, they would also remove the difference in the number of hits due to other variables such as different sensor systems.

Section 7

SUMMARY AND CONCLUSIONS

An observer's performance in a target acquisition task may be a function of properties of the target and imagery, (image resolution, grazing angle, degree of camouflage) and the operator's decision rule (presumed to be related to the payoff for correctly identifying a target and the consequences associated with false alarms).

The separation of factors that influence the observer's attitudes from those that influence sensitivity is a major contribution of the psychophysical application of statistical decision theory. TSD techniques allow for studying the effects of a particular variable, whether the effects of the variable are on detectability or on the location of the observer's criterion. The experimenter can thus observe whether systematic changes in the value of the variable result in different points along a single ROC curve (indicating that the variable affects the location of the observer's criterion) or points located on different ROC curves (indicating that the variable affects target detectability or observer sensitivity). Values of d' and ß can be calculated for various experimental conditions. Manipulation of an independent variable may result in changes in d', ß, or both.

It has been experimentally demonstrated that d', the TSD measure of visual sensitivity, is not contaminated by the effects of variables which shift the observer's response criterion. Further, d' values, unlike the different threshold values obtained through the use of the various classical psychophysical methods, remain relatively invariant when measured by different experimental procedures. When the observers were required to say "yes" or "no" in response to a designated time interval that sometimes contained a signal, d' estimates were found to approximate those obtained when the observer was required to rate his confidence that a signal was or was not present. According to Coombs (1970),

> "It is, of course, a very substantial accomplishment for a theory to provide predictability and integration over a wide variety of experimental conditions and procedures. It appears that for a given observer and a given S/N ratio, d' is reasonable constant over variations in ß induced by changing the prior odds and payoff matrix and, for the most part, over variations in procedures..." (pg 199)

TSD offers significant advantages over classical operator performance paradigms. It is certainly valuable in the evaluation of human target acquisition performance in complex sensor systems which may consist of the sensor, automatic target recognizer (ATR), display, and human observer. In such a system, the ATR has associated hit and false alarm rates which have been adjusted to some predetermined criterion level. If the human observer views only those images which the ATR has declared as target images, the overall system false alarm rate should be somewhat reduced. However there is also the possibility that the human will reject actual targets declared by the ATR. The observer's decision criterion plays an important role in overall system performance. The theory of signal detection allows researchers to quantify the operator's decision criterion as well as his visual sensitivity.

REFERENCES

Coombs, C. H., Dawes, R. M., and Tversky, A. (1970) Mathematical psychology, an elementary introduction. Prentice-Hall, Inc.

Egan, James P. (1975) Signal detection theory and ROC analysis. New York: Academic Press.

Gescheider, George A. (1976) Psychophysics method and theory. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Green, David M. and Swets, John A. (1966) Signal detection theory and psychophysics. New York: John Wiley and Sons, 1966.

Greening, C. P. (1986). Targeting decisions with multiple data sources: a status report. (NWC TP 6706). China Lake, Ca.: Naval Weapons Center.

Greening, Charles P. and Foyle, David C. (1988). The man/machine interface in multisensor targeting: program status and preliminary guidelines. (NWC TP 6864). China Lake, Ca.: Naval Weapons Center.

Kuperman, Gilbert G., Wilson, Denise L., and Perez, William A. (1988). Relocatable target acquisition performance with simulated synthetic aperture radar imagery. (AAMRL-TR-88-025). Wright-Patterson AFB, Ohio: Armstrong Aerospace Medical Research Laboratory.

Ozkaptan, Halim. (1979). Evaluation of the utility of the theory of signal detection for target acquisition studies. (Technical Report 405). U.S. Army Research Institute for the Behavioral and Social Sciences.

Pollack, I. and Norman, D.A. (1964) A non-parametric analysis of recognition experiments. Psychonomic Science, 1, 125 - 126.

Skolnik, Merrill I. (1980) Introduction to radar systems. New York: MacGraw-Hill.

Smith, M., and Wilson, E. A. (1953) A Model for the auditory threshold and its Application to the multiple observer. Psychological Monogr., 67, No.9.

Swets, J. A. (1964) Signal detection and recognition by human observers: contemporary readings. New York: John Wiley and Sons.

Swets, John A., Tanner, Wilson P. Jr., and Birdsall, Theodore G., (1964) Decision processes in perception In John A. Swets (ED)

Signal detection and recognition by human observers. New York: John Wiley and Sons.

Tanner, W. P. Jr., and Swets, J. A. (1954) A decision-making theory of visual detection. _Psychological Review_, _61_, 401 - 409.

GLOSSARY

a posteriori probability	Probability of the truth of each hypothesis after an event e_k has occurred
a priori probability	Probability of any particular state of the world prior to an observation
ATR	Automatic Target Recognizer
β	(Beta) Operator decision criterion
CR	Correct rejection
δm	(Delta-m) A nonparametric statistic
d'	(d-prime) Measure of perceptual sensitivity
e_k	information or evidence about h_i
FA	False alarm
h_i	ith possible state of the world
H_j	jth hypothesized state of the world
$l_{ij}(e_k)$	(Likelihood ratio) Ratio of the probabilities of an event under two different hypotheses
N	Noise distribution
$p(e_k \mid h_i)$	Probability of e_k conditional upon h_i
ROC	Receiver operating characteristic
Σ	(Sigma) Summation
SN	Signal plus noise distribution
TSD	Theory of Signal Detection

U.S. Government Printing Office